REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
12-03-2001	Final	

4. TITLE AND SUBTITLE

Microstereopsis

5a. CONTRACT NUMBER
F6256200M9081

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Mr Hiroshi Mataki

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Kansei Research Institute
17 Chudoji Minami-machi
Shimogyo-ku, Kyoto 600-8813
Japan

8. PERFORMING ORGANIZATION REPORT NUMBER

N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AOARD
UNIT 45002
APO AP 96337-5002

10. SPONSOR/MONITOR'S ACRONYM(S)

AOARD

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AOARD-004001

12. DISTRIBUTUION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13 SUPPLEMENTARY NOTES

14. ABSTRACT
This report results from a contract tasking Kansei Research Institute as follows: The contractor will investigate 3D viewing by microstereopsis (1) To verify that microstereopsis can generate effective 3D images; (2) To evaluate its effectiveness in terms of depth perception and physiological fatigue.

15. SUBJECT TERMS

Stereopsis

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18.NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE		25	Terence J. Lyons, M.D.
U	U	U	UU		**19b. TELEPHONE NUMBER** (Include area code) +81-3-5410-4409

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

FINAL REPORT

TO THE AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Contract # F6256200M9081 (AOARD-00-01)

Research on 3D viewing by Microstereopsis

March 8, 2001

Hiroshi Mataki
Kansai Research Institute (KRI)

Mel Siegel
The Robotics Institute
Carnegie Mellon University

Masaaki Takeuchi
Kansai Research Institute (KRI)

KRI

CONTENTS

ACKNOWLEDGEMENT

ABSTRACTS

Conventional 3D systems using two plano-stereo images would give rise to physiological fatigue known as "virtual reality sickness" due to accommodation-convergence conflict, and other conflicts among depth cues. In order to solve such a problem we proposed a new concept of stereopsis; **microstereopsis**. The basic idea of microstereopsis consist of the following hypothesis:

(1) Human visual system is capable of creating stereo images even when camera interocular separations, and thus on-screen disparities, are far smaller than human interocular separation.

(2) When "center-of-interest compensation" is added to the small on-screen disparities, the depth perception will be enhanced effectively.

(3) Crosstalk between left eye and right eye can be perceived as foreground and background blur, which (in contrast to its usual perception as ghosting) is acceptable to the human eye-brain system.

For the purpose of verifying the hypothesis of microstereopsis, we assembled a display system which can provide plano-stereo images with a disparity far smaller than the human interocular distance. Then, we conducted experiments to see (i) what interocular distance is sufficient to have depth-perception, and (ii) what crosstalk is acceptable as blur. Finally, we have obtained the following results:

(1) A stereo pair corresponding to as little as 1 mm of interocular distance separation can reliably be perceived as a stereo image.

(2) Even when the crosstalk is 10%, one can definitely perceive stereo without ghosting whenever the interocular distance of a stereo pair is less than or equal to about 5 mm. When the crosstalk is as much as 40%, interocular separation of 2 and 3 mm still stimulate stereo perception without substantial perception of ghosting.

From an engineering point of view, it can therefore be expected that zoneless autostereoscopic displays are available using some devices on the market such as an LCD or a new display equipped with switchable louver filter.

2

1. Introduction

Since 1997, KRI and Dr. Mel Siegel have been co-working on the study of stereopsis and have published the hypothesis, **microstereopsis**, for the first time in the SPIE (The International Society for Optical Engineering) Conference held on January 25, 1999[1]. The basic concept of microstereopsis, in comparison with conventional stereopsis, can be summarized as follows:

(1) In conventional stereopsis, on-screen disparities usually correspond to human interocular separation. For this, the conflict between convergence and accommodation will cause physiological fatigue.

(2) Even when on-screen disparities are far smaller than inter-ocular separation, human visual system is still capable of creating stereo images. When "center-of-interest compensation" is added to the small on-screen disparities, the depth perception will be enhanced effectively.

(3) In conventional stereopsis, in which on-screen disparities are large, crosstalk between left eye and right eye is perceived as ghosting. This will not be acceptable to naked human eyes. On the other hand, however, when on-screen disparities are small, crosstalk between left eye and right eye is perceived as foreground and background blur, which is acceptable.

We can think, therefore, that the concept briefly described above suggests the following engineering concepts:

(1) We no longer need complete left eye/right eye separation when displaying plano-stereo images. Rather, we could use a display projecting two plano-stereo images with small directional luminescence disparities, each of which corresponds to left eye and right eye, respectively.

(2) Because the on-screen disparities are allowed to be small compared with inter-ocular separation, we may not necessarily need to use two cameras or a two-lens camera. This could be practically realized using known single-lens multiple-aperture stereo methods.

These engineering concepts are thought to suggest that zoneless autostereoscopic

3

displays without any eyewear can be obtained. Also, it is expected that such an autostereoscopic display can be observed from any direction, unlike some of the conventional stereo displays such as the Lenticular-type display.

For the purpose of quantitatively obtaining basic understanding of phenomenological features of microstereopsis, we submitted a research proposal to AOARD on November 8, 1999. The goal of the proposed research can be summarized as follows:

(1) To verify the hypothesis of microstereopsis by assembling a display system which is to provide us plano-stereo images with a disparity far smaller than the human interocular distance.

(2) To evaluate the effectiveness of microstereopsis upon depth perception in stereo vision in the light of physiological fatigue.

In order to verify that microstereopsis is effective enough to create stereo images, we attempted to evaluate the effect of variation of on-screen disparities and the degree of center-of interest compensation on depth perception. For this approach, we have assembled prototype hardware, which can provide plano-stereo images with a disparity far smaller than human interocular distance. Then, we have conducted pilot experiments to evaluate the effectiveness of microstereopsis in the light of depth-perception.

In this paper, we will report the research on microstereopsis, which has been supported by AIR FORCE OFFICE OF SCIENTIFIC RESEARCH. First, the detail of the hardware, by which we conducted the experiments, will be described in the relation with the basic concept of our research. The experiments and their results will be described in Section 3. In Section 4, discussion on the experimental results and the future prospect for zoneless autostereoscopic display systems will be made. Because Hiroshi Mataki and Mel Siegel shared the preparation of this report, some parts will be overlapped.

2. 3D-Stereoscopic Workstation: Hardware, Application Programming Environment, Prototype Applications, and Outcome of Pre-Pilot Experiments

(1)Objective

The primary objective of this section is to document the hardware and software infrastructure and the application programming software environment of a 3D-stereoscopic computer graphics and video workstation. Its secondary objective is to document a suite of prototype applications designed to evaluate several human factors issues in the presentation and perception of 3D-stereoscopic imagery. Its tertiary objective is to present the results of several pre-pilot experiments that were conducted using the prototype applications, and, based on the outcomes of these experiments, to suggest directions and experimental designs for future human factors research in 3D-stereoscopic perception.

(2)Background

The human eye-brain system perceives depth by processing the differences, called disparities, between its left and right eye perspectives on the world. Artificial 3D-stereoscopic displays stimulate depth perception by steering separate 2D ("flat") right and left eye perspective views to the corresponding eyes.[1]

The left and right eye views may be collected with a stereoscopic still or video camera pair that captures physical images from different perspectives, they may be generated by a computer graphics engine that simulates a physical camera pair viewing a 3D world model, or they may be generated by a hybrid system in which, first, a physical image pair is interpreted to build an internal 3D world model, and, second, the desired left and right eye perspectives are calculated by the graphics engine.

The job of a 3D-stereosocopic display system is then to present each viewer's eyes with the appropriate perspective views. Assuming for now that the appropriate (viewer

---

[1] This report considers only 3D-stereoscopic displays in which depth perception is stimulated by left-right pairs of "flat" (2D) images. Specifically, it ignores "volumetric" displays, e.g., systems in which the part of a scene that belongs at a particular depth is projected onto a cyclically-moving screen at instants when the screen is at the appropriate depth.

dependent) perspectives are known and the 3D world model is sufficiently complete that the required views can be generated, the engineering task is simply to steer one flat image to the left eye and the other flat image to the right eye. Systems for achieving this end can usefully be divided into two broad classes: spatial multiplexing and temporal multiplexing. In spatial multiplexing systems an optical system directs imagery drawn on each part (e.g., pixel) of the display screen into one or the other eye as is appropriate. In temporal multiplexing systems a shuttering system cyclically occludes or exposes one or the other eye, and corresponding left and right eye imagery is drawn on the screen synchronously.

Temporal multiplexing systems are generally preferred over spatial multiplexing systems because spatial multiplexing compromises image resolution. In spatial multiplexing systems, the resolution seen by each eye is the total number of display pixels divided by the number of perspective views rendered. The number of perspective views rendered is at least two. It is correspondingly worse if individual viewers viewing the same screen are each shown different perspectives corresponding to their different locations with respect to the screen.

On the other hand, temporal multiplexing systems are not without engineering challenges. In particular, they must be capable of switching between perspective views sufficiently rapidly that flicker perception is suppressed. While this requirement does not place particularly challenging demands on the electronics, both existing CRT phosphors and existing LCD materials have in practice been substantially optimized for 25-30 frames per second interlaced display or 50-60 frames per second progressive displays. When these displays are run at even the minimum required factor of two faster than this, unacceptable left-right channel crosstalk is observed.

An acceptable compromise is to use a moderately high end computer graphics workstation equipped with a monitor capable of 90 or more non-interlaced frames per second, and to limit the number of perspective views to the two that are geometrically correct for the primary viewer looking at the screen straight on. In practice, the perspective distortion seen by secondary viewers' off-axis is easily accommodated.

Time multiplexing requires cyclic shuttering between the display screen and the eyes with a 180 degree phase difference between the left and right eye sequences, i.e., each must be closed whenever the other is open. To achieve this it is necessary for the viewer to wear some sort of spectacles or goggles. These may be active, e.g., with mechanical or (preferably) electro-optical shuttering in front of each eye, or passive,

e.g., with orthogonal polarizing filters at the eyes and a polarization-switching device at the screen. The prototype system uses passive left and right hand circular polarizers at the eyes and an on-screen device that correspondingly switches polarization synchronously with the temporal sequencing of left and right eye frames.

One of the tentative conclusions of the "pre-pilot" human factors experiments performed to date with the prototype system is that most viewers possess a conveniently accessible perceptual region in which left-right eye disparity is large enough to stimulate depth perception but small enough that crosstalk between left and right eye temporal channels is perceived as a slight blur in the scene's foreground and background. This blur is unobjectionable in comparison with the ghosting that is perceived when disparities are larger. This tentative observation suggests the possibility, in the future, of developing hybrid 3D-stereoscopic display systems that have simultaneously temporal and spatial multiplexing elements. These hybrid systems could be used autostereoscopically (without glasses) while being free of the annoying "viewing zones" characteristic of spatially multiplexed autostereoscopic systems.

(3)Hardware Implementation

Platform

The workstation platform is a Silicon Graphics Incorporated (SGI) 320 NT personal computer running the Windows NT 4.0 Service Pack 5 operating system. It is equipped with dual 650 MHz Pentium III processors. This workstation's display memory is organized with "quad buffering" to support 3D-stereoscopic display of video and computer graphics.

Ordinary computer graphics workstations provide dual buffers: the content of the "front" buffer is displayed dynamically while the contents of the "back" buffer is being overwritten by the graphics engine or (in principle) by the video stream directly. When the back buffer has been completely updated, pointers to the front and back buffers are interchanged: the former "back" buffer, now the front buffer, is displayed statically, while the former "front" buffer, now the back buffer, is updated. The sequence is repeated cyclically, typically at 60 Hz.

Dual buffered systems can display 3D-stereoscopic imagery only by employing some sort of additional "trick", e.g., by drawing left and right images anamorphically

compressed vertically in an above-below format and injecting an "extra" vertical synchronization pulse mid-frame. In contrast, quad buffered systems have left and right, front and back buffers. The left and right front buffers are displayed sequentially and cyclically, while left and right back buffers are overwritten in the background. Either (left or right) back buffer can be swapped with the corresponding (left or right) front buffer at the beginning of the display cycle following completion of its update.

The workstation hardware furthermore provides an externally accessible signal whose state indicates whether a left buffer or a right buffer is being displayed. This signal is sent to external hardware that controls the polarization-switching hardware.

The workstation is also equipped with a single channel video digitizer capable of streaming video to the screen via applications support libraries such as Windows Multimedia Viewer and Apple QuickTime. Live 3D-stereoscopic video can be achieved with the aid of external hardware, for example, the StereoGraphics Corporation's "View/Record" product, which multiplexes asynchronous video streams from two sources (e.g., cameras) anamorphically compressed horizontally into a side-by-side format. The side-by-side format can, in software, be decoded, missing columns replicated or interpolated, and the result displayed as time-multiplexed stereo video.

This functionality has been verified on the workstation, but the additional hardware and software are not part of the delivered package. The reason is that there is a much better way to achieve this capability with the SGI 320 NT's higher-end counterpart, the SGI 540 NT. This version supports dual channel digital video input. This being a much higher quality, straightforward, and future-oriented (i.e., HDTV-compatible) route to achieving live video, it was deemed diversionary to pursue the single channel side-by-side option in depth.

The manufacturer's summary specification, manuals in PDF format, and driver updates can be obtained at http://support.sgi.com/nt/product/320/.

Temporal Multiplexing

There are two main options for temporal multiplexing of left-right image channels: an "eye-side" option that uses shuttering eyewear, e.g., LCD shutters synchronized with the display by wire or by infrared signal, and a "display-side" option using switched polarization at the display and passive polarizing spectacles at the eyes. The display-side option is more comfortable, i.e., the spectacles are similar to lightweight ordinary

sunglasses, whereas in contrast the LCD shutter goggles are heavy and uncomfortable. The display-side option is also not at all annoying in terms of interference between multiple monitors (3D-stereoscopic or ordinary) in the same room, whereas the eye-side option is extremely annoying in this respect. The display-side shuttering is more expensive than the shuttering eyewear, but tolerably so; furthermore, when there are multiple viewers, the display-side option rapidly becomes the less expensive technology per viewer, since additional polarizing spectacles are very inexpensive. The prototype workstation thus uses the display-side shuttering option.

The actual instantiation uses the StereoGraphics Corporation "Z-Screen" product. This is a so-called "π-cell" at the screen and left- and right-hand circular polarizing spectacles at the eyes.

Circular polarization is produced by a combination of linear polarization and birefringence. The birefringent material is oriented with its fast axis rotated 45 degrees ($\pi/4$ radians) with respect to the linear polarization direction; the thickness of the birefringent material is whatever is required to produce a 90 degree ($\pi/2$ radian) phase shift between the linear polarization components in the fast-axis and slow-axis directions. This produces either left- or right-hand circular polarization, depending on the sign of the 45 degree offset between the linear polarizer and the fast axis of the birefringent material. An additional 180 degree (π radian) phase shift complements the handedness of the circular polarization: left to right, or right to left, as the case may be. In practice, of course, one would not build a passive circular polarizer of one handedness and reverse it with an electro-optically active 180 degree phase shift, but would rather use simply an electro-optically active device that switches between plus 90 degrees and minus 90 degrees, hence the name "π-cell" to describe the full range.

The Z-Screen product consists of an anti-reflection coated screen that mounts securely but removably on the monitor bezel, e.g., with mating Velcro "hook-and-loop" adhesive strips, and a small interface box that converts the state signal from the workstation into the appropriate voltage amplitude and waveform to power the Z-Screen. The interface box has a three-position toggle switch that selects normal operation, off, and reversed operation for situations where the sign of the control signal may be the reverse of the assumed convention.

The manufacturer's summary specification sheet is available at http://www.stereographics.com/, click on "products", and scroll to the Z-Screen summary, click on "more information".

(4)Software

Programming Environment

3D-stereoscopic demonstrations and applications are programmed for the Windows NT operation system using Microsoft Visual C/C++ and the OpenGL graphics library. The OpenGL library supports 3D-stereoscopic applications that use a left/right/front/back quad buffering paradigm. With the "stereo option" hardware with which the workstation is equipped, Silicon Graphics provides a CDROM with the drivers required to appropriately access the quad buffered hardware.

The general structure of a 3D-stereoscopic program follows this template:

1. #include " gl (backwards/)gl.h"

   #define STEREO_ENABLED

   Also: link to OpenGL32.lib at compile time.

2. EnableStereoWindow()

   ShowWindow()

   glGetBooleanv (GL_STEREO, &Test)

   Examine "Test" for correct response from hardware, exit if incorrect.

3. glDrawBuffer (GL_BACK)

   Setup for drawing in the left and right back buffers. If drawing graphics that may not overwrite every pixel, at this point it would be appropriate to call glClearColor() to define a black or dark gray background color and glClear() to actually fill the background. Dark gray is usually preferable to black, since it helps suppress perception of left-right channel crosstalk.

4. glDrawBuffer (GL_BACK_LEFT)

   Enable writing to the back left buffer, then write left view there ...

   glDrawBuffer (GL_BACK_RIGHT)

   Enable writing to the back right buffer, then write right view there.

5. Toggle pointers to front and back buffers so newly written back buffer becomes front buffer and old front buffer becomes new back buffer. Newly written left and right eye views are displayed sequentially beginning with next full display cycle.

6. Absent any user input to do something else, loop back to step (3).

There are, of course, alternative approaches to accessing the workstation's graphics engine and display hardware. These are generally "low level" approaches that bypass

10

the high level functionality provided by the OpenGL library; as such, they may be expected to achieve improved run-time performance, primarily execution speed, at the expense of a much more burdensome programming task. Since the purpose of this workstation is to allow and encourage rapid prototyping of a wide variety of 3D-stereoscopic demonstrations and human factors experiments, minimizing program development time and maximizing the researchers' opportunities to explore alternatives are more important goals than is maximizing run-time speed. Thus the high level approach is in general to be preferred. Nevertheless, it is important to recognize that there may be future circumstances, e.g., if it is decided to investigate human factors relating to display of 3D-stereoscopic HDTV imagery, when it will be necessary to employ a low level access to the hardware in order to achieve the required performance.

Demonstration Programs and Prototype Experiments

3D-Stereoscopic Workstation Functionality Check

StereoGraphics Corporation provides, with the Z-Screen, source code (main program and auxiliary files) and an executable file for a simple test program, RedBlue, that, as per the above paradigm, creates a stereo window, clears the background to dark gray, draws a red rectangle in the middle of the left buffer, draws a blue rectangle in the same position (no disparity) in the right buffer, and displays the result. If all the hardware and its software support is working correctly the viewer sees a magenta (red + blue) rectangle in the screen plane. If any component is not working correctly then what is seen is diagnostic for the problem, for example, a red rectangle seen in the left eye and only the gray background seen in the right eye suggest that the right back buffer is not being written correctly.

We have extended this program to provide several additional functionality checks and demonstrations of basic system functionality and elementary 3D-stereoscopic effects.

Depth Perception vs. Disparity

The historical origin of this project can be traced to KRI's early interest in what has come to be called "kinder gentler stereo", i.e., an approach to 3D-stereoscopic image generation and display that overcomes the stressful artifacts, collectively termed "virtual reality sickness", that are associated with conventional "geometrically correct stereo".

In the course of research aimed at characterizing and quantifying the human factors issues, the concept of "microstereopsis" was developed collaboratively by CMU and KRI. By microstereopsis is meant (1) a 3D-stereoscopic image generation technique in which the "geometrically correct stereo" paradigm is modified by reducing the interocular separation to the minimum value that stimulates adequate depth perception, and (2) a 3D-stereoscopic image display technique in which left and right views are shifted so as to reduce their disparity to near-zero at the center-of-interest. The second statement is equivalent to saying that the scene is shifted forward or backward as necessary to put the center-of-interest at the screen plane. As a practical matter, in an integrated system, particular one that relies heavily on computer graphics technology for image generation or capture, processing, rendering, and display, the distinction between what is done on the image generation side and what is done on the image display side becomes pretty fuzzy. Thus in practice the conditions for achieving microstereopsis can be achieved in part or in whole by optical means at image capture time, and in part or in whole by algorithmic means at image processing, rendering, and display time.

Early in this study it was observed that adequate depth perception for scene appreciation and understanding is stimulated by 3D-stereoscopic image pairs for which the effective interocular separation is surprisingly small: under favorable circumstance, as little as 1-2% of the normal human interocular separation. Hence the term "microstereopsis" for this regime.

To quantify the effect, an experiment was designed - and software to prototype it was written - in which viewers were presented with side-by-side randomly constructed and selected pairs of 3D-stereoscopic images (some of which were actually "flat") and asked "which one has more depth". This experiment was actually first programmed and pre-piloted using an older SGI Indy Unix workstation. However it is of sufficiently crucial importance that it is appropriate to port it to the demonstration platform.

The structure and tentative outcome of the pre-pilot experiments that were conducted has been reported in a refereed journal article, Reference [2], and in more detail elsewhere in this report, so only the above brief summary is provide here. Figure 1 (below, left) illustrates the screen presented to the subjects, and Figure 2 (below, right) illustrates the outcome of the pre-pilot experiments.
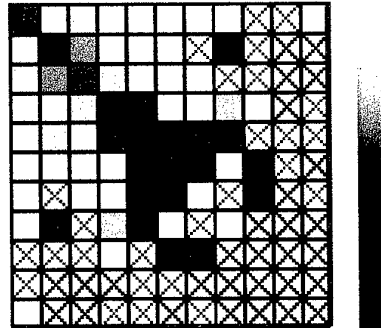
click the pic that "has more depth"
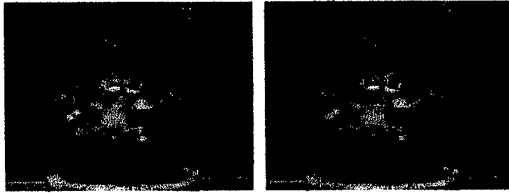
Figure 1(left): Example of choice presented to subjects in pre-pilot relative depth perception experiment.

Figure 2 (right): Left to right: interocular separation in one picture. Top-to-bottom: interocular separation in other picture. Gray scale: percent (black=0, white=100) of times the side with more depth is correctly identified. Average of ten informal subjects, result symmetrized about the diagonal.

Crosstalk Perception vs. Disparity

When a temporally multiplexed 3D-stereoscopic display is viewed directly, i.e., *without* the requisite demultiplexing eyewear, the left and right eyes both see the left- and right-eye's images, apparently simultaneously. Because the two images are substantially identical, i.e., they differ only by their perspective disparity, the visual appearance is similar to the visual appearance of the familiar offset-doubled TV image that occurs when there are multipath or impedance mismatch problems at the receiver. In both cases the visual effect is called "ghosting". Ghosting is also observed in temporally multiplexed 3D-stereoscopic displays when the channel separation is imperfect, i.e., when there is *crosstalk* between the channels due to engineering imperfections in the demultiplexing system. Some common causes of crosstalk in 3D-stereoscopic display systems are, with shutter goggles, inadequate occlusion of the "off" eye because the LCD shutters allow too much light leakage in their nominally off state, and, with Z-Screen shutters, inadequate occlusion of the "off" eye because the Z-screen does not produce perfectly circularly polarized light – it is rather elliptically polarized – and/or the nominally circularly polarizing eyeglasses ("analyzers") fail to completely block light of the "wrong" polarization handedness. Whatever the technical cause of the

13

crosstalk, its perception as ghosting is objectionable. Indeed, it is so objectionable, so inimical to the effective and pleasant perception of stereopsis, that elimination of crosstalk becomes probably the primary objective of 3D-stereocopic display system designers.

Whereas it is normally taken for granted that the physical phenomenon of *crosstalk* and the perceptual phenomenon of *ghosting* are one-and-the-same effect, careful open-minded consideration suggests that crosstalk may not always be perceived as ghosting. It seems plausible, indeed, it seems obvious, that if disparity is small enough then crosstalk may be observed as *blurring* rather than doubling of the image. The case of microstereopsis seems particularly sanguine in this respect, in that its disparity is zero at and around the scene's center-of-interest, so any perception of blurring would be confined to background and foreground regions. Since background and foreground regions are *normally* somewhat blurred by the depth-of-focus effect, crosstalk with microstereopsis may be perceived as quite natural, hence unobjectionable. In fact this hypothesis is borne out by even casual observation: when a temporally multiplexed 3D-microstereoscopic display is observed without demultiplexing eyewear the imagery usually looks entirely natural. An unprepared viewer is typically unaware that he or she is looking at overlapping left-eye and right-eye images on a stereoscopic display.

Viewing a temporally multiplexed 3D-stereoscopic display without demultiplexing eyewear is, of course, the worst case for crosstalk: both eyes see both images without attenuation. If total crosstalk is unobjectionable, then surely partial crosstalk will also be unobjectionable. Furthermore, if total crosstalk is borderline objectionable, then partial crosstalk may well be completely unobjectionable. This possibility is an intriguing one, for it raises the hope that, with microstereopsis, the stereoscopic system designer's primary constraint, the elimination of crosstalk between left- and right-eye channels, could be substantially relaxed. That this is actually the case is tentatively demonstrated by pre-pilot experiments that are described in the next few paragraphs.

Initial informal experiments were conducted, prior to the beginning of this project, using a hardware approach: the control and power electronics of a wired pair of LCD shutter glasses was modified to controllably degrade their effectiveness in occluding each channel in its nominally-off state. The results were encouraging, but quantitation was clumsy due to the nonlinearity of the relationship between the control signal and resulting occlusion. The next step, undertaken during the course of the present project, was a software solution: crosstalk would be synthesized algorithmically by mixing a

fraction of the right-eye image into the left-eye view and vice versa.

The software that accomplishes this is straightforward. The desired crosstalk is specified by a single parameter, C, whose range is conveniently defined to be 0 to 0.5. In the generic stereoscopic display algorithm the 4th step:

glDrawBuffer (GL_BACK_LEFT)

Enable writing to the back left buffer, then write left view there ...

glDrawBuffer (GL_BACK_RIGHT)

Enable writing to the back right buffer, then write right view there.

is generalized to depend on the crosstalk parameter C as follows:

glDrawBuffer (GL_BACK_LEFT)

Enable writing to the back left buffer

modified_left_view = (1.0-C)*left_view + C*right_view

write modified left view to left back buffer

glDrawBuffer (GL_BACK_RIGHT)

Enable writing to the back right buffer

modified_right_view = C*left_view + (1-C)*right_view

write modified right view to right back buffer

When C=0 the resulting imagery has no crosstalk – except, of course, for the inevitable display hardware crosstalk – and when C=0.5 the resulting imagery is the same – except for the brightness attenuation due to the eyewear – as what would be seen if the display were viewed without eyewear.

Armed with this algorithm and its implementation within the general 3D-stereoscopic imagery display program, pre-pilot experiments were conducted in which several subjects were shown, in random order, using the Z-Screen display and circularly polarizing demultiplexing eyeglasses, a set of 3D-stereoscopic image pairs with interocular separations in the range 0 to 10 mm and crosstalk parameter C in the range 0.0 to 0.5. Each subject was asked to classify each image in one of four classes: {{flat, stereo},{ghosted, normal}}. "flat" means no depth was perceived, "stereo" means there was at least some perception of depth; "ghosted" means there was noticeable doubling of the image in the foreground and background regions, "normal" means the foreground and background regions looked normal, i.e., either sharp or unobjectionably blurred, as if because of the depth-of-field effect.

The hypothesis is that there is a perceptual region in the (disparity, crosstalk) plane in which it is simultaneously true that (i) disparity is large enough to stimulate depth

perception AND (ii) crosstalk is small enough that is perceived unobjectionably as background and foreground blur rather than objectionably as ghosting. As shown in Figure 3 in the following section, the tentative conclusion of these pre-pilot experiments confirms this hypothesis. These preliminary experiments are also discussed in References [3] and [4], and Figure 3 appears in Reference [2].

(5)Future Research Opportunities

Formalizing the Pre-Pilot Experiments on Microstereopsis and Crosstalk Perception

In the course of building the described 3D-stereoscopic display workstation, two "pre-pilot" experiments were conducted and described: (1) perception of depth as a function of small interocular separation (i.e., disparity) with disparity adjusted to zero at the center-of-interest (i.e., "microstereopsis"), and (2) delineation of the perceptual region in which interocular separation is large enough to stimulate depth perception and crosstalk is small enough to be perceived as unobjectionable blur. The hypothesis of (1) is that very small on-screen disparities, corresponding to interocular separations as small as a few percent of the normal human interocular separation, are sufficient to convey a sense of depth, and that furthermore the depth sense so conveyed is "kinder and gentler", i.e., negligibly prone to cause "virtual reality sickness" in comparison "geometrically correct stereo".

The hypothesis of (2) is that there is a perceptual region in which disparity is large enough to stimulate depth and crosstalk is small enough to be perceived as unobjectionable blur, and that furthermore this region can be exploited by 3D-stereoscopic display system designers by allowing them to relax the heretofore strict requirement to minimize crosstalk at almost any cost. The pre-pilot experiments tentatively confirm both hypotheses. It is recommended that future research be undertaken to formalize these experiments, using strict human factors experimental protocols and obtaining data with a statistically valid sample population size and makeup.

16

References [2], [3], and [4] suggest the additional hypothesis that relaxing the heretofore assumed "fundamental" requirement to absolutely minimize crosstalk in 3D-stereoscopic displays may enable the practical engineering of a new class of displays that would be "autostereoscopic", i.e., viewable in stereo without special eyewear, and "zoneless", in contrast to currently available autostereoscopic displays, all of which must perforce divide the viewing space into sharply delineated angular zones, and requiring the viewers' head to be located such that each eye is located in an appropriate left- or right-eye zone. The references suggest several possible engineering implementations of this concept based, e.g., on a switched electro-optical active counterpart of the passive "privacy screens" used on ATM machines, laptop computer displays, etc. Several concepts for realizing this screen, e.g., suspended particle technology and LCD technology, suggest themselves as plausible possibilities. It is recommended that future research be undertaken to identify additional realization possibilities, to characterize them with respect to anticipated suitability and development cost, to prototype a working example, and to design and perform human factors experiments to test the hypothesis. If correct, prospects for profitable commercialization seem enormous, inasmuch as the primary barriers to the widespread utilization of 3D-stereoscopic displays appear to be the requirement for eyewear in temporally multiplexed systems and the inevitability of viewing zones in all current spatially (angularly) multiplexed autostereoscopic systems. This point is also discussed in Section4 in the relation with possible solutions.

3. Experiments and Results

3-1. Ability of depth perception

First, we conducted an experiment in order to derive the relation between on-screen disparity and the depth-feeling of viewers when they see pairs of pictures displayed side-by-side. [See Fig.1 in Section 2]   In this experiment, we presented each viewer with two stereo images and asked him or her to indicate whether the left picture, neither picture, or the right picture seemed to have more depth.   The subject was given no additional guidance as to what "having more depth" means.   The pictures were randomly chosen from image data set by a computer program that picked images only on-the-fly from a set of 41 that were taken 1 mm apart at a working distance of about 30cm with a 12.5mm focal length lens.   Each view was shown 80 pairs of stereo images in two sets of 40.

In the first 40, one image was always flat, i.e., the left-eye and right-eye sub-images were actually identical.   In the second 40 both were nominally stereo, although by chance sometimes one or both might have zero disparity.   Within each 40 there were 5 sequences 8 x 2 images.   Each of the 8 was drawn from a random Poisson weighted disparity distribution.   The first 8 had a mean camera interocular separation of 4.5 mm, the second 8 had a mean of 3.5 mm, and so on down to the fifth 8 with a mean of 0.5 mm.   The flat images in the firs 40 were taken from perspectives randomly displaced from the midpoint of the data set by the same statistical distribution that randomizes the disparities.   The subjects were told that the task "starts easy and gets more difficult as it progresses", but they were not told that in the first sets one of the pictures was always flat.

The experimental result is summarized and shown in Fig. 2.   From this result, it can easily be seen that across the sample population, a stereo pair corresponding to as little as 1 mm of interocular separation can reliably be distinguished from an image that is actually flat.

18

3-2 Perception of crosstalk

"Crosstalk" and "ghosting" are often used as if they are synonymous. In reality, however, they are not: "crosstalk" is the electrical or optical mixing of left- and right-eye image channels, whereas "ghosting" is a particular mode of perception of imagery that has been degraded by crosstalk. It is clear therefore that other modes of perception are possible; for example, crosstalk may be so small as to be imperceptible. Even when crosstalk is as large as it possibly can be, as when an eyewear-based 3D-stereoscopic display is viewed without eyewear, the crosstalk may still be imperceptible, e.g., if the on-screen disparities are very small.

As the second experiment, we attempted to see how a pair of images could be perceived when the crosstalk is as large as it possibly can be, and the on-screen disparity is increased smoothly from zero to the point where the crosstalk is first perceived. Also, we attempted to clarify if it would be perceived as *ghosting*, or is it rather — our hypothesis — perceived only as degrading of the spatial resolution, i.e., as *blur*.

We initially conducted a simple test of our hypothesis by modifying a conventional LCD shutter-glasses controller to give it adjustable crosstalk. The protocol was to show a subject a stereopair taken with normal interocular separation, and to instruct him or her to increase the crosstalk adjustment until ghosting became perceptible. At this crosstalk level, the subject was then shown image pairs with reduced interocular separation. The experiment was repeated with three informally recruited subjects, leading to the preliminary conclusion was that there is indeed a region in which crosstalk is relatively large, interocular separation is relatively small, and in that region crosstalk is perceived as blurring rather than ghosting and scene depth is perceived accurately and comfortably.

To confirm quantitatively the existence of a region in which on-screen disparity is large enough to stimulate depth perception and small enough to be perceived as blur rather than ghosting, and, if it exists, to map its borders, a program was written to display stereopairs with variable interocular separation and variable crosstalk. A stereopair with a selected interocular separation, e.g., 5 mm, is randomly selected from a subset of the image data set (the above-described 41 images collected with 12.5 mm lens, 30 mm

object distance, and mechanical center-of-interest compensation), a crosstalk percentage (e.g., 0% -> the baseline crosstalk inherent in the display hardware, 25% -> 1:3 mixture of left image in right eye view and vice versa, 50% -> left eye view and right eye view are the same mix of left eye image and right eye image, 100% -> completely pseudoscopic except for the baseline crosstalk inherent in the display hardware). A subject is asked to place each image in one of two *stereo perception class*es, {~stereo | stereo}, and one of two *crosstalk perception class*es, {~ghosting | ghosting}, where the tilde (~) signifies "not". For each classification a point is plotted with a symbol that labels which of the four combined classes {~stereo, ~ghosting|}, etc. The results was summarized for three informally selected subjects in Figure 3.

Perceptions of Depth and Ghosting as a function of Interocular Separation and Crosstalk
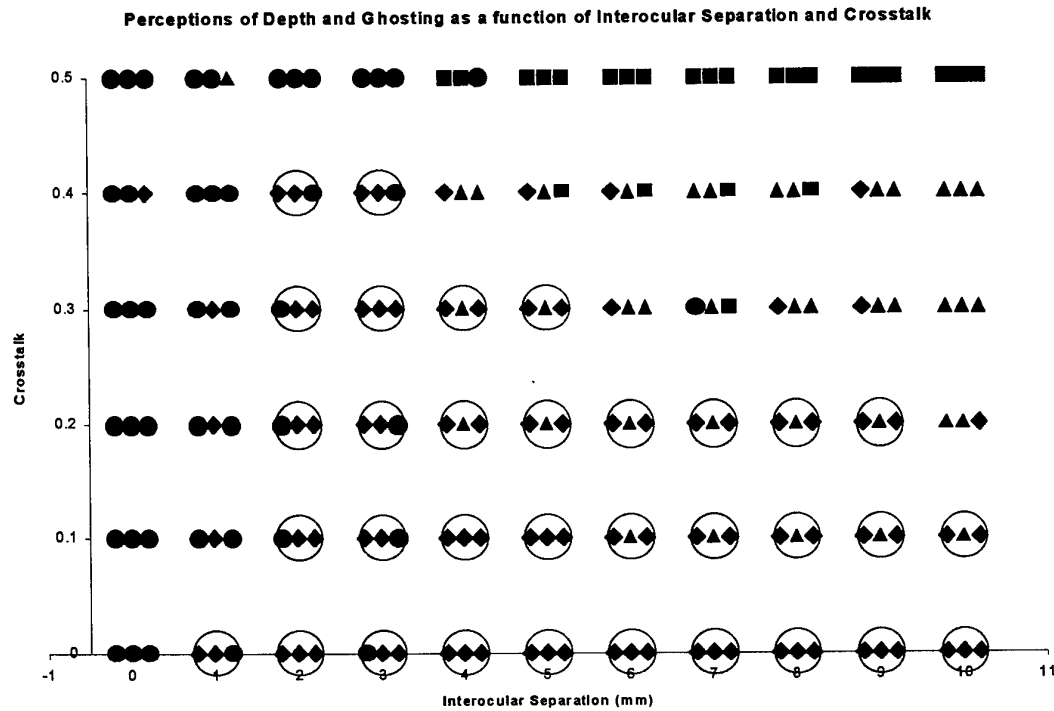
Figure 3: Perceptions of stereo and crosstalk. Stereopairs are generated with randomly selected interocular separation (0 to 10 mm) and crosstalk parameter C (see text for definition). Subjects are asked to classify each displayed stereopair in one of four classes corresponding to perception of the image as being flat or in stereo and ghosted or not ghosted. Legend: {stereo, ~ghosting}->DIAMOND, {stereo, ghosting}->TRIANGLE, {~stereo, ~ghosting}->CIRCLE, {~stereo, ghosting}->SQUARE. LARGE OPEN CIRCLES around a group of three small symbols denotes a group in which two or three out of three subjects agree that the image is in stereo and without ghosting. Note the bell shaped region of LARGE OPEN CIRCLES wherein this desirable condition is realized. This region corresponds to the perception of stereo without the perception of ghosting despite the presence of crosstalk. When the crosstalk parameter C is 0.1 (i.e., the pixel-by-pixel blending of 90% the intensity of the "correct" image with 10% of the intensity of the "wrong" image), stereo without ghosting is definitively perceived whenever the interocular separation is less than or equal to about 5 mm. When the crosstalk parameter C is as much as 0.4 (i.e., the left eye is shown 60% left eye image and 40% right eye image), interocular separations of 2 and 3 mm still stimulate stereo perception without substantial perception of ghosting. Note the reasonableness of this shape: as expected, even when the interocular separation is very small indeed, i.e., 1 mm, stereo perception is suppressed by sufficient crosstalk, even though when the interocular separation is this small the crosstalk is perceived as slight foreground and background blur rather than as ghosting.

21

4. Discussion

The experimental results that have been shown in the previous section indicates that there is a perceptual region in which disparity is small enough that — in the presence of crosstalk — it is perceived as blur, yet it is nevertheless large enough to stimulate binocular stereopsis.

From an engineering point of view, this suggests the possibility zoneless autostereoscopic displays. That is, zoneless displays would be achievable because of the tolerability of crosstalk between left and right eye channels. Under these circumstances it is not necessary to completely separate the left-eye and right-eye channels; it is adequate that an appropriate bias be created, such that the left eye "sees more" (e.g., via an illumination disparity) of the left-eye image than it sees of the right-eye image, and vice versa. This kind of bias can be achieved by means of a suitably non-lambertian screen over an otherwise conventional display. In a time-multiplexed system the bias (e.g., illumination gradient) would be made to alternate directions synchronously with the alternation of left-eye and right-eye images on the display. Practical implementations could be achieved with electronically switched louver filters based on, e.g., suspended particle displays, LCDs, and other technologies.

At this time we can describe only in a general sort of way, without yet being able to give numerical values for physical or psychophysical parameters, as these will have to be determined by future experiments, how we would go about engineering this system. One approach is simply to illuminate the display screen (if it is transmissive, e.g., an LCD) or filter it (if it is emissive, e.g., a CRT) in an angular pattern that looks brighter or dimmer depending on the viewing angle, i.e., it is "non-lambertian" in a particularly designed sort of way. The two eyes thereby see different screen brightnesses corresponding to the different azimuths from which they view the screen. Passive screens with precisely this property (but with stronger gradients than we probably want) are actually commercially available, e.g., 3M's "Privacy Shield" material for bank ATMs and laptop computer screens (e.g., for frequent flyers who want their screens to be invisible from the adjacent seat). This screen material is a microfabricated "venetian blind"; the generic device is called a louver filter, the concept of which is illustrated in Figure 4.

Figure 4: Two states (R, L) of louver filter. Looking up from this caption toward the filter, in the R state the bias favors the right eye, in the L state it favors the left eye.

Two engineering challenges remain to be overcome to turn this idea into a practical microstereoscopic display: (1) we need an electronically switchable louver filter, and (2) the gradient needs to be strong enough between the eyes that sufficient bias is achieved, but not so strong over the full range of viewing azimuth that the illumination difference between the two states is annoying. Depending on the outcome of measurement of the psychophysical factors, this tradeoff may limit the display's useful range of viewing angles. On the other hand, even in the worst case the idea should nevertheless be workable for viewing the display more-or-less head-on, and even in that worst case it should still be far less restrictive about head position, in both azimuth and range, and than are the lenticular and barrier displays currently on the market.

We suggest, for example, that an electronic louver filter can be implemented using suspended particle display technology 0 as illustrated schematically in Figure 5. It uses opaque dielectric particles with permanent dipole moments suspended in a transparent dielectric liquid. The particles are oriented as desired by an electric field produced by electrodes patterned on the windows. This technology is currently in pilot production for "smart windows" for automatic control of indoor sunlight. It

seems to require only more complex electrode patterning and driver electronics to make it into an electronic louver filter. Other possibilities, e.g., liquid crystal display based electroholographic devices, and also under consideration.
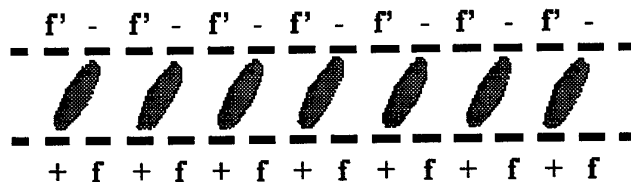

Figure 5: Electrode pattern and polarization required to produce the L state of Figure. Transparent electrodes are held at positive (+) or negative (-) voltage, or allowed to float (f,f'). Elongated opaque suspended particles in liquid carry a permanent dipole moment. Interchanging + and f on the lower electrode, - and f' on the upper electrode changes device to the R state. Note wave-like character of activation.

23

5. Conclusion

In order (1)to verify the hypothesis of microstereopsis, and (2) to evaluate the effectiveness of microstereopsis upon depth perception in stereo vision in the light of physiological fatigue, we assembled a display system which can provide plano-stereo images with a disparity far smaller than the human interocular distance. Then, we conducted experiments to see (i)what interocular distance is sufficient to have depth-feeling, and (ii)what crosstalk is acceptable as blur. As a result, we can conclude as follows:

   (1) A stereo pair corresponding to as little as 1 mm of interocular distance
       separation can reliably be perceived as a stereo image.
   (2) Even when the crosstalk is 10%, one can definitely perceive stereo without
       ghosting whenever the interocular distance of a stereo pair is less than or equal
       to about 5 mm. When the crosstalk is as much as 40%, interocular separation
       of 2 and 3 mm still stimulate stereo perception without substantial perception of
       ghosting.

These results suggest that it is not necessary to completely separate the left-eye and right-eye images to create appropriate stereo images as have been done in the conventional 3D display systems. From an engineering point of view, this can be said to suggest that zoneless autostereoscopic displays are available using some devices on the market such as an LCD or a new display equipped with switchable louver filter.

References

1. Mel Siegel, Yoshikazu Tobinaga, Takeo Akiya; "Kinder Gentler Stereo", *Proceedings of SPIE*, Vol.3639 (1999)

2. Siegel, M. and S. Nagata, *Just Enough Reality: Comfortable 3D Viewing via Microstereopsis.* IEEE Transactions on Circuits and Systems for Video Technology, 2000. **10**(03): p. 387-396.

3. Siegel, M. *Just Enough Reality: Microstereopsis, and the Prospect of Zoneless Autostereoscopic Displays.* International Conference on Image Processing (ICIP) 2000. September 2000, Vancouver BC Canada.

4. Siegel, M. *Perceptions of Crosstalk and the Possibility of a Zoneless Autostereoscopic Display.* SPIE/IS&T Stereoscopic Displays and Applications. 2001. San Jose CA: SPIE.